

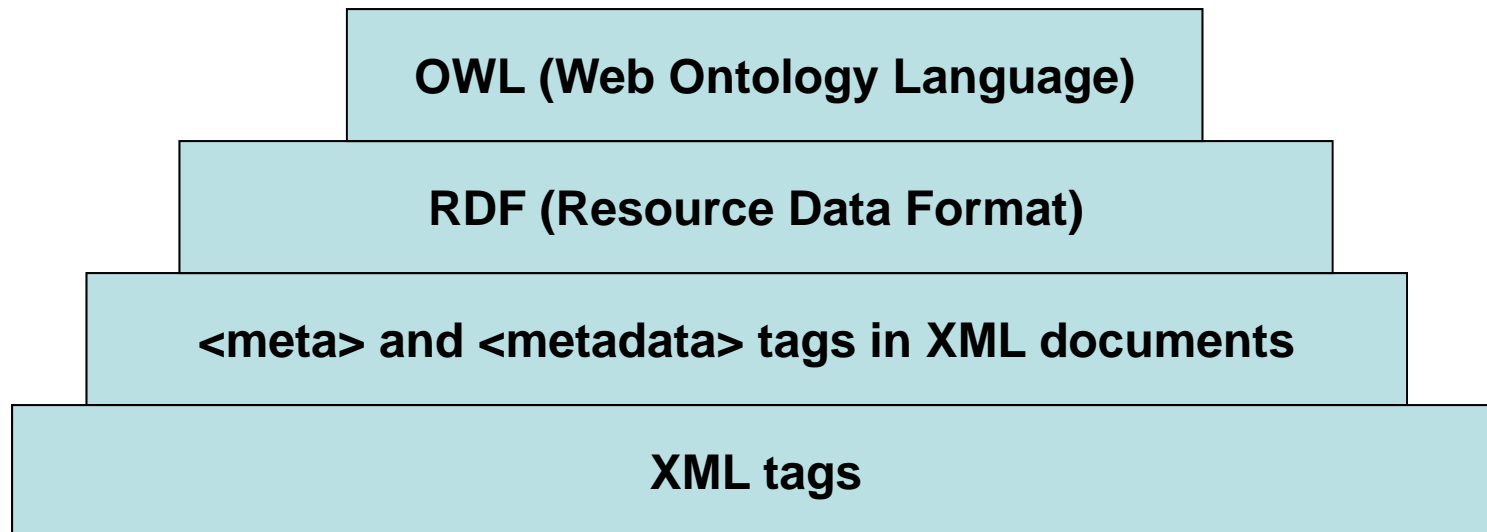
A103 Search Using Metadata

James A. Larson
Larson Technical Services
and Information Today, Inc

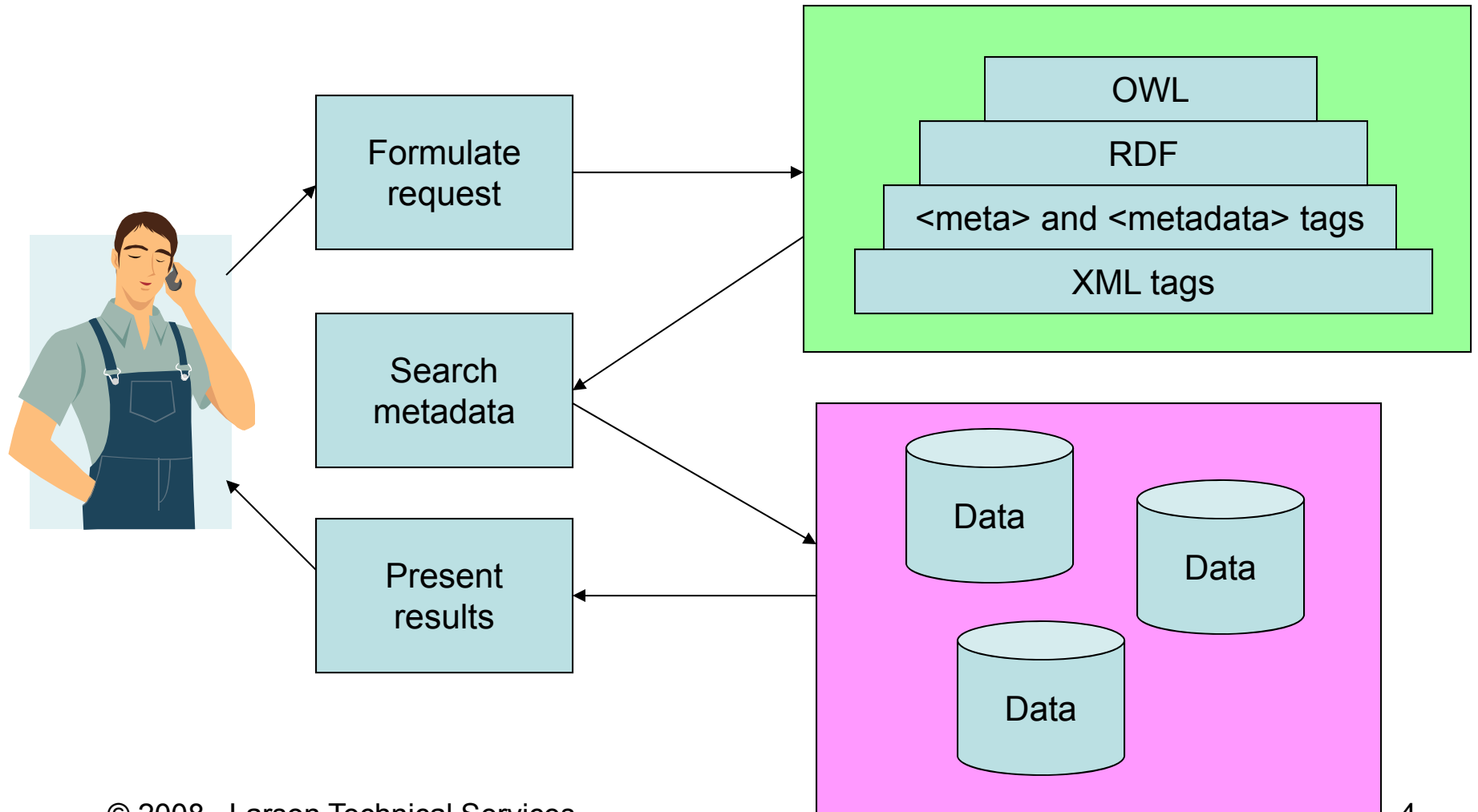
Metadata

- Metadata = data about data
- Used to locate requested data quickly
- Examples of metadata from real life
 - Table of contents at the front of a book
 - Index at the back of a book
 - Card catalog in a library
 - Directory of the day's events posted in a hotel lobby

Standard XML-based metadata



The Big Picture of Voice and XML



XML Tags

- Used to locate information of the desired type

```
<employee>  
  <name> Smith </name>  
  <salary> 22000 </salary>  
  <hireDate>2006-04-09</hireDate>  
</employee>
```

OWL (Web Ontology Language)

RDF (Resource Data Format)

<meta> and <metadata> tags in XML documents

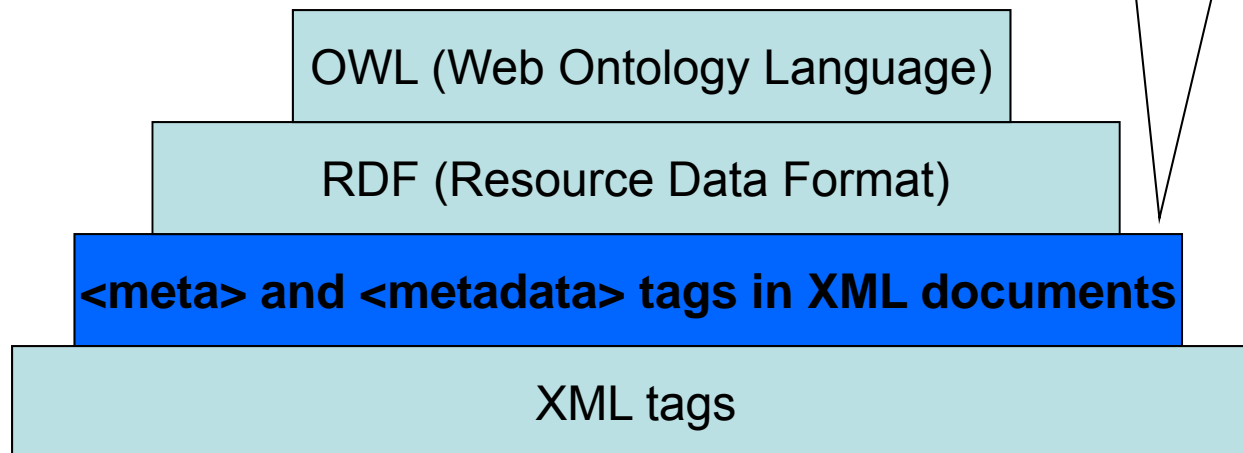
XML tags

<meta> Tags

Used to

- Provide documentation
- Search <meta> tag contents instead of searching ALL of the words in a document

```
<meta name="keywords"  
content="tree, plant, perennial,  
trunk, branches" />
```



<metadata> Tags with “Dublin Core” Annotations

```
<metadata>
  <rdf:RDF
    xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:rdfs = "http://www.w3.org/TR/1999/PR-rdf-schema-19990303#"
    xmlns:dc = "http://purl.org/metadata/dublin_core#">
  <rdf:Description about="http://www.example.com/meta.vxml"
    dc>Title="Directory Enquiry Service"
    dc:Description="Directory Enquiry Service for London in VoiceXML"
    dc:Publisher="W3C"
    dc:Language="en"
    dc>Date="2002-02-12"
    dc:Rights="Copyright 2002 John Smith"
    dc:Format="application/voicexml+xml" >
    <dc:Creator>
      <rdf:Seq ID="CreatorsAlphabeticalBySurname">
        <rdf:li>Jackie Crystal</rdf:li>
        <rdf:li>William Lee</rdf:li>
      </rdf:Seq>
    </dc:Creator>
  </rdf:Description>
</rdf:RDF>
</metadata>
```

- Used to search and query

OWL (Web Ontology Language)

RDF (Resource Data Format)

<meta> and <metadata> tags in XML

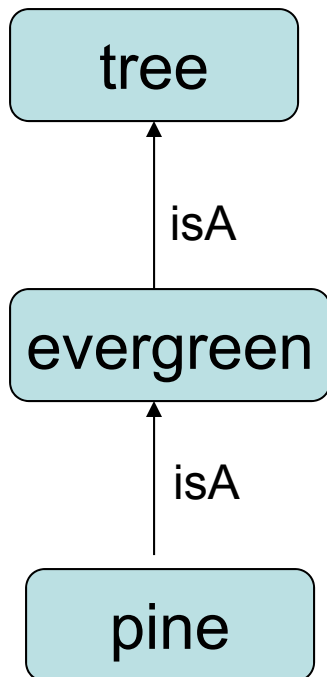
XML tags

What Is RDF?

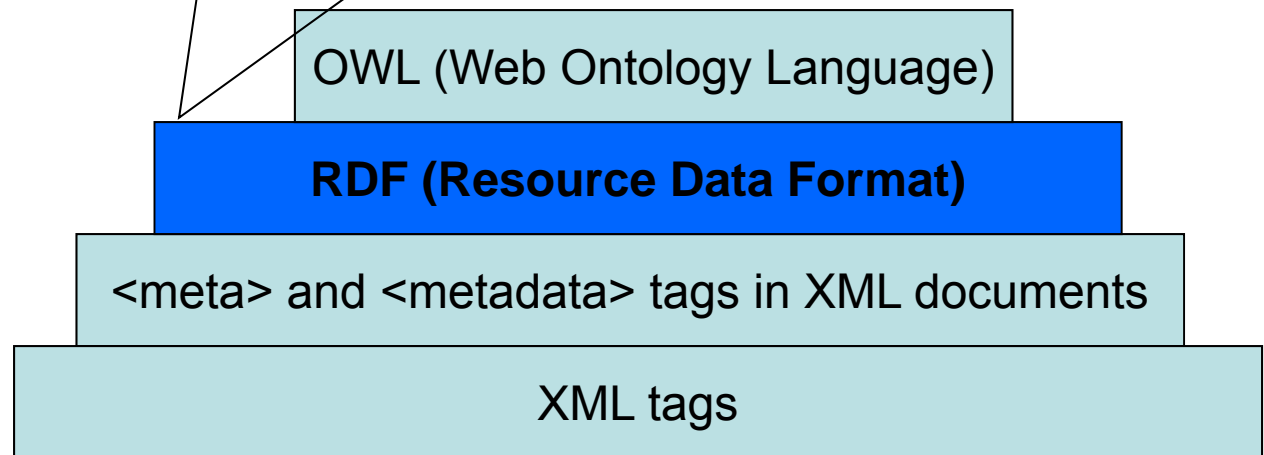
- RDF stands for Resource Description Framework
- RDF is a framework for describing resources on the web
- RDF provides a model for data, and a syntax so independent parties can exchange and use code
- RDF is designed to be read and understood by computers
- RDF is not designed for being displayed to people.
- RDF is written in XML.
- RDF is a part of the W3C's Semantic Web Activity.
- RDF is a W3C Recommendation.

RDF (Resource Description Framework)

- Used primarily for link chasing during browsing

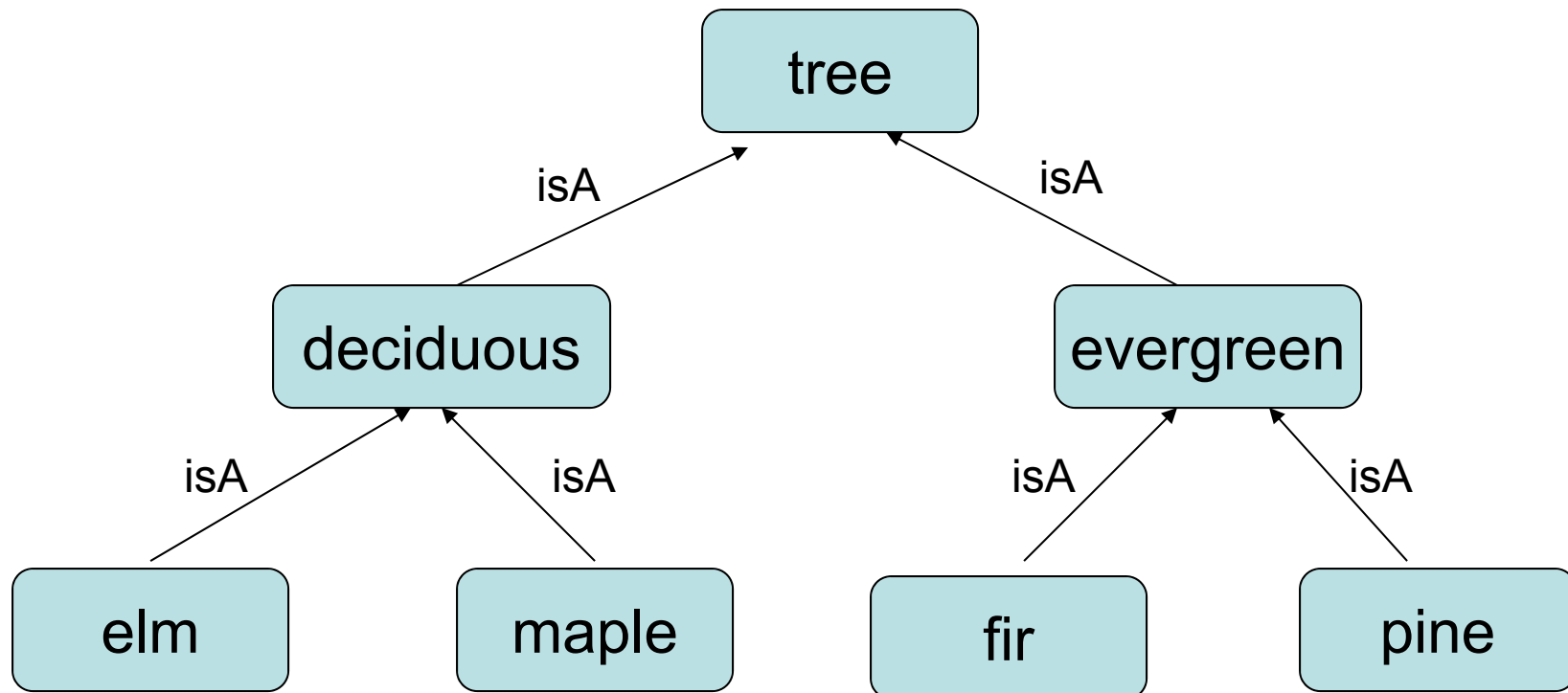


```
<Description about="some.uri/evergreen">
  <isA resource="some.uri/tree"/>
</Description>
<Description about="some.uri/pine">
  <isA resource = "some.uri/evergreen " />
</Description>
```



What Is Ontology?

- Ontology is a description of things and their relationships.



What Is OWL?

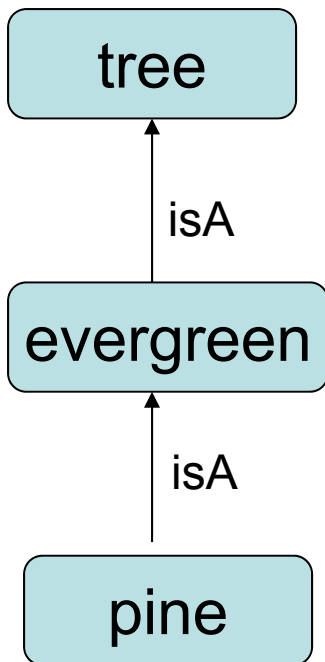
- OWL stands for Web Ontology Language.
- OWL is built on top of RDF.
- OWL is for processing information on the web.
- OWL was designed to be interpreted by computers.
- OWL was not designed for being read by people.
- OWL is written in XML.
- OWL has three sublanguages.
- OWL is a web standard.

OWL Is Different from RDF

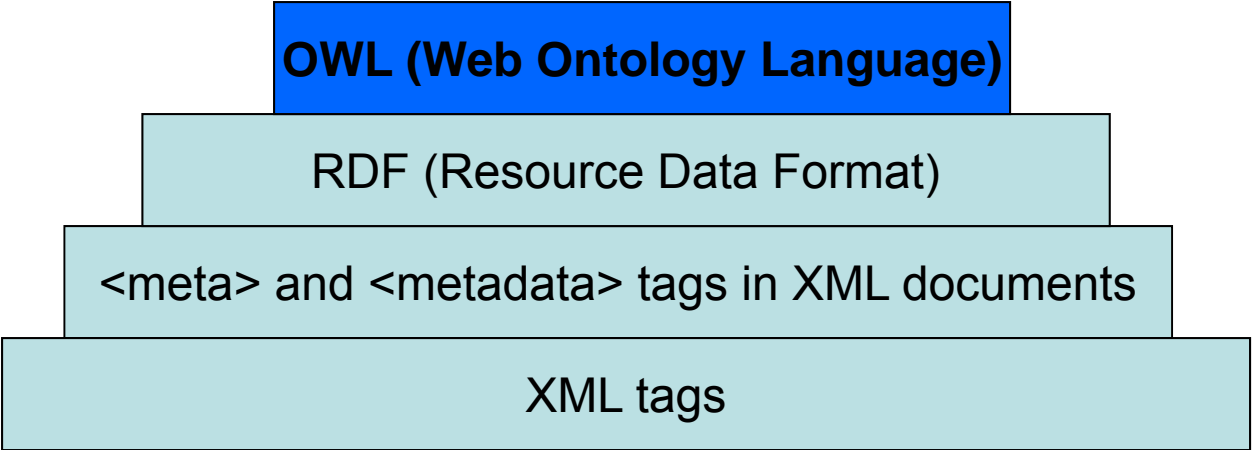
- OWL and RDF are much of the same thing, but OWL is a stronger language with greater machine interpretability than RDF.
- OWL comes with a larger vocabulary and stronger syntax than RDF.

OWL (Resource Description Framework)

- Used primarily for reasoning

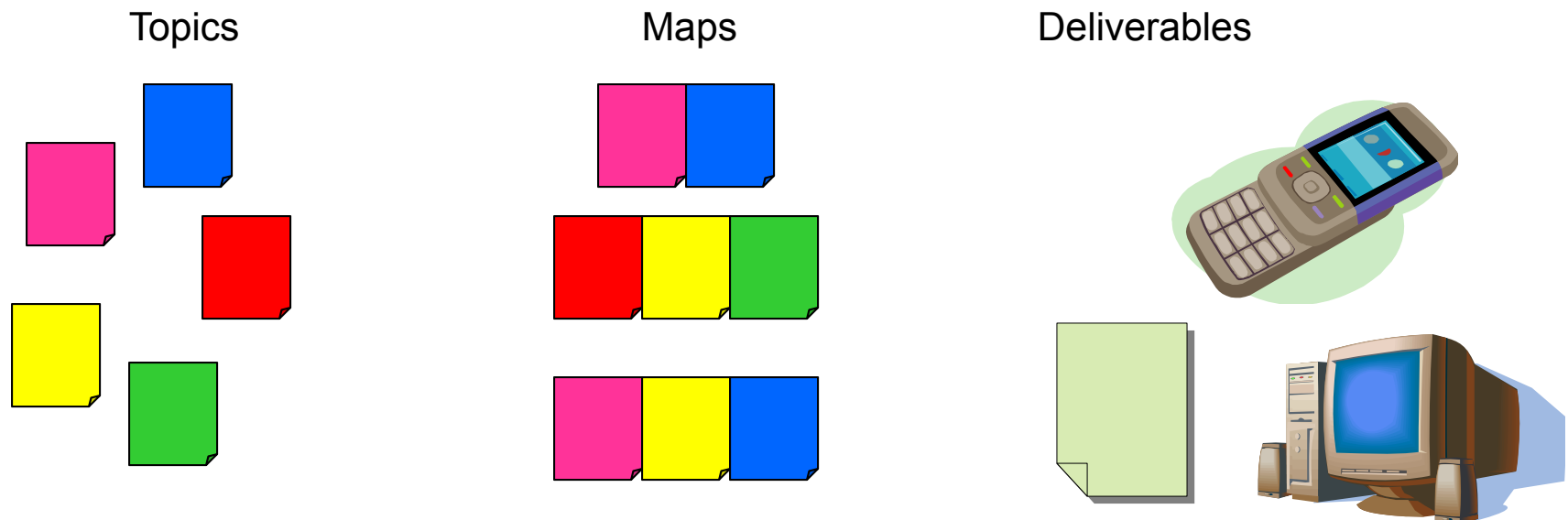


pine isA evergreen \wedge evergreen isA tree \rightarrow
Pine isA tree



What Is DITA?

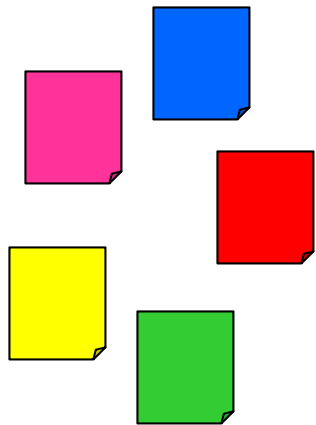
- Darwin Information Typing Architecture
- An XML-based architecture for authoring, producing, and delivering technical information.



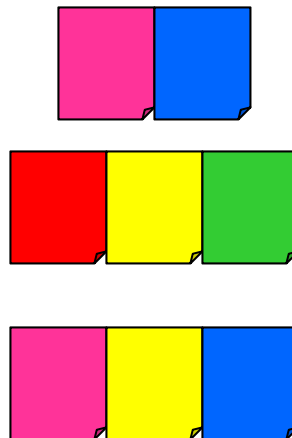
Using DITA

- Search topics
- Combine and format topics
- Present to user

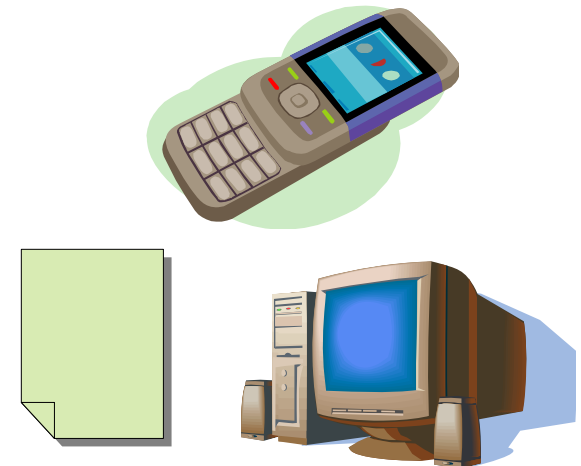
Topics



Maps

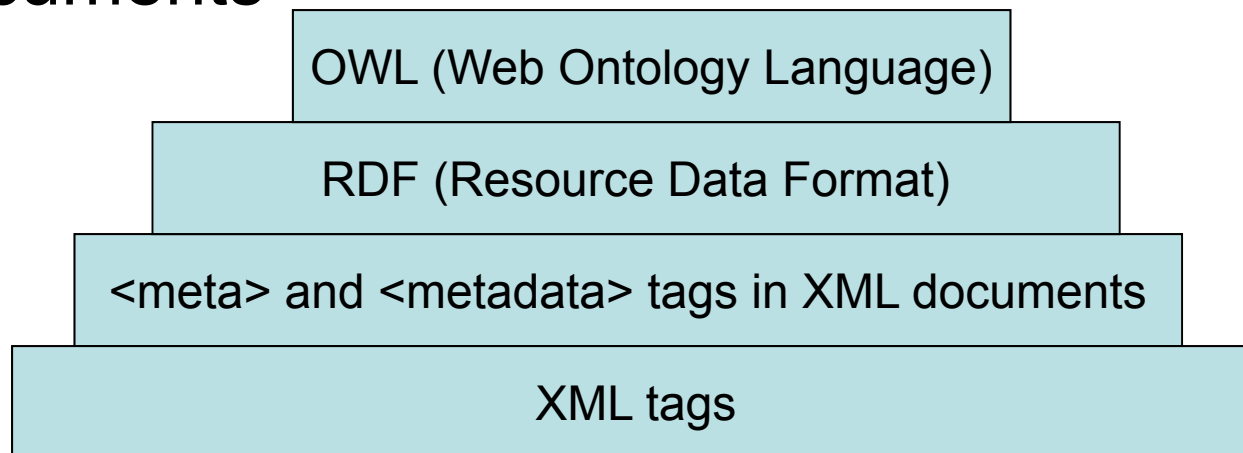


Deliverables



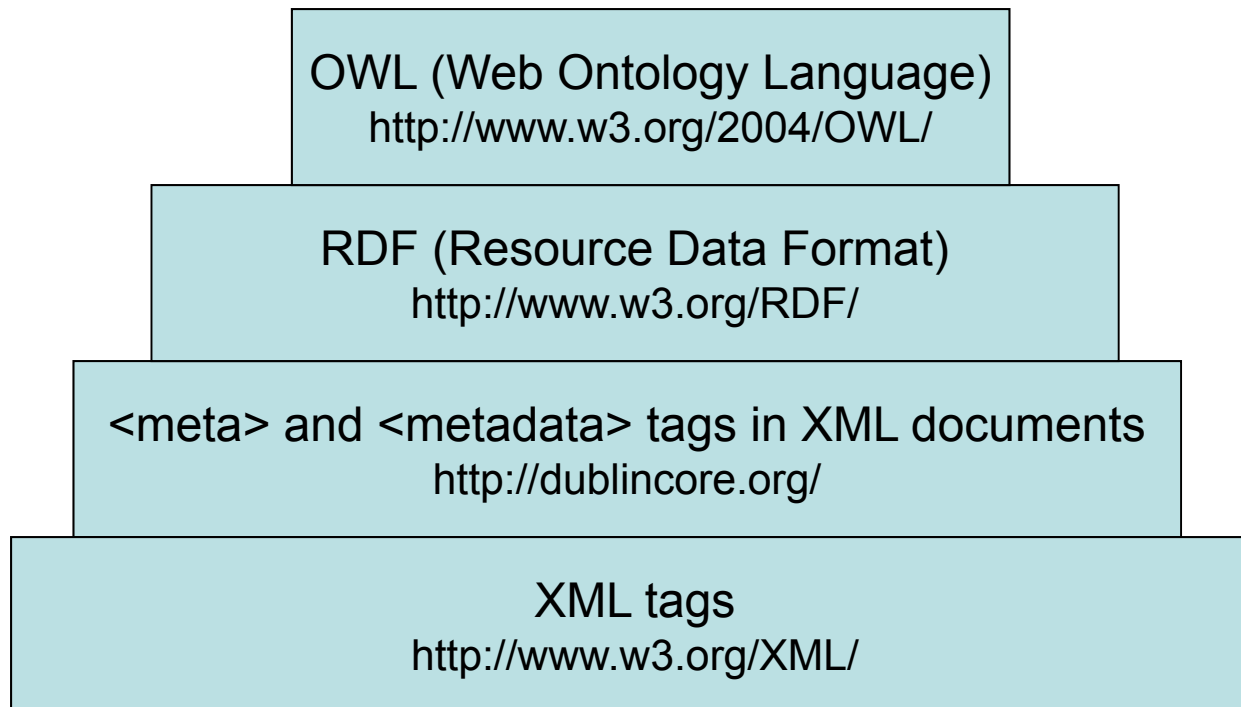
Where Does Metadata Come From?

- Reuse existing metadata
- Build it yourself
 - Manual development
 - Automatic extraction of metadata from documents



Summary

- Use metadata to reflect the content of your web pages
- Make metadata part of your searching strategy
- Review detailed language specifications



Questions

